UNIVERSITI PUTRA MALAYSIA

DEVELOPMENT OF AN ASSERTION MODEL OF INTEGRITY CONSTRAINTS IN OBJECT-ORIENTED DATABASES

BELAL MOHAMMAD ZAQAİBEH

FSKTM 2006 11

DEVELOPMENT OF AN ASSERTION MODEL OF INTEGRITY CONSTRAINTS IN OBJECT–ORIENTED DATABASES

By

BELAL MOHAMMAD ZAQAIBEH

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

June 2006

DEDICATION

To the memory of my Grandfather,

To my Parents: Mohammad and Ne'mat,

To my Wife: Maisa and my Son: Mohammad,

To my Brothers and Sisters.

Belal

Abstract of thesis presented to Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

DEVELOPMENT OF AN ASSERTION MODEL OF INTEGRITY CONSTRAINTS IN OBJECT-ORIENTED DATABASES

By

BELAL MOHAMMAD ZAQAIBEH

June 2006

Chairman: Associate Professor Hamidah Ibrahim, PhD

Faculty: Computer Science and Information Technology

Object-Oriented Databases (OODBs) have been designed to support large and complex programming projects. The data accuracy, consistency, and integrity in OODBs are extremely important for developers and users. Checking the integrity constraints in OODBs is a fundamental problem in database design. Existing OODB Management Systems (OODBMSs) lack to a capability of an ad-hoc declarative specification of enforcing and maintaining integrity constraints that are appeared among attributes in association, composition, and inheritance hierarchies' relationships.

A critical problem in the existing OODBs is that they cannot support User-Defined Constraints (UDCs) that can be defined in classes with composition (logical or physical composition) and inherence (single or multiple inheritance) hierarchies. Integrity constraints in the current

OODBMSs are maintained either by disallowing and rolling back transaction or modifying operations that may produce a violation. The constraints must be maintained in the backward direction along the class composition hierarchy as well as in the forward direction.

In this work an Assertion Model of Integrity Constraints (AMIC) is proposed. The AMIC keeps the derivation path along with the attributes' relationships that are derived from association, composition, and inheritance hierarchies. The AMIC techniques are designed to implement the needed functions that are collecting the attributes' relationships and checking the integrity constraints. Moreover, AMIC keeps UDCs with their relationships in both single classes and multilevel classes (intra-class and inter-class). Furthermore, the AMIC can maintain constraints in a single object and a set of distributed objects (intra-object and inter-object). Therefore, this makes the new model extendable and can be integrated with any existing constraints' service.

A new technique called Detection Method (DM) is designed to check the Object Meta Data (OMD) to detect the constraints violation before it occurs. The AMIC is designed for both Centralized Integrity Maintenance (CIM) and Application-Oriented Integrity Maintenance (AOIM). The AMIC can also enforce and maintain structural and logical integrity constraints, in addition to enforce and maintain redundant, inconsistent, and duplicate constraints.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PEMBANGUNAN MODEL PENERAPAN KEKANGAN INTEGRITI DALAM PANGKALAN DATA BERORIENTASI OBJEK

Oleh

BELAL MOHAMMAD ZAQAIBEH

Jun 2006

Pengerusi: Profesor Madya Hamidah Ibrahim, PhD

Fakulti: Sains Komputer dan Teknologi Maklumat

Sistem pengkalan data berorientasikan objek (OODB) direka untuk menyokong projek pemprograman yang kompleks dan berskalar besar. Ciri-ciri seperti keetepatan data, konsistansi pangkalan data serta kewibawaan data dalam OODB amat penting bagi pengguna dan pembangun perisian. Penyemakan kewibawaan dalam OODB merupakan dilema asas yang dihadapi oleh pereka pangkalan data. Ini disebabkan sifat terwujud OODB yang tidak berupaya untuk memproses spesifikasi deklaratif yang bertujuan mengekalkan kewibawaan konstren-konstren yang ada kaitan dengan hierarki persekutuan, komposisi dan pewarisan. Sifat sedia ada OODB inilah yang menyebabkan ialah ia tidak berupaya untuk meyokong kekangan atau konstren yang tertakrif dalam komposisi sesebuah

kelas (komposisi secara logik mahupun fizikal) ataupun pewarisan (tunggal mahupun berganda) dengan sebaiknya.

Buat masa ini, konstren kewibawaan dalam OODB disokong melalui penggulungan balik sesuatu transaksi, tidak membenarkan transaksi tersebut ataupun melaui pengubahsuaian operasi tersebut. Cara-cara sedemikian besar kemungkinan boleh menyebabkan keadaan pangkalan data menjadi tidak konsisten kerana pemeliharaan konstren dilakukan dalam dua arah bertentangan (hadapan dan belakang) di dalam hierarki komposisi kelas.

Dalam tesis ini, Model Penerapan Kewibawaan Konstren (AMIC) dicadangkan untuk meyokong konstren takrifan pengguna. AMIC mengekalkan laluan penerbitan serta sifat perhubungan yang diperoleh dari hierarki persekutuan, komposisi dan pewarisan. AMIC digunakan untuk melaksana fungsi-fungsi yang mengumpul sifat perhubungan serta memeriksa kewibawaan konstren. AMIC juga mengekalkan konstren serta hubungan mereka dalam kelas tunggal dan berganda (antara dan intra kelas). Selebihnya, AMIC juga dapat mengekalkan hubungan sedia ada di dalam sesebuah objek dan juga antara objek-objek tertabur. Justeru itu, prestasi akan bertambah baik dan membolehkan model ini diluaskan dan disepadukan dengan sebarang perkhidmatan konstren yang sedia ada.

Teknik baru yang dinamakan Kaedah Pengesanan (DM) digunakan untuk mengenalpasti Objek Meta Data (OMD) bagi tujuan mengesan pencabulan terhadap pangkalan data sebelum ia berlaku. AMIC di cipta untuk menyokong kewibawaan terpusat dan juga kewibawaan berorientasikan aplikasi. AMIC juga dapat mengekalkan kewibawaan struktur dan logik untuk model-model yang berasaskan OODB. AMIC juga dapat mengekalkan konstren yang lewah, tak konsisten dan salinan.

ACKNOWLEDGEMENTS

In the name of ALLAH, I heartiest would like to thank my supervisor Associate Professor Dr. Hamidah Ibrahim, for her incredible guidance, continuous support, and encouragement. Always having time for me and readily providing her technical expertise. I owe more than I can ever repay, her valuable directions and suggestions are very helpful in my research. Only has the successful completion of this work become possible due to her supervision, she is the first person to thank for making my PhD. program at the Universiti Putra Malaysia a very enjoyable experience. Her high stand of diplomatic power and professionalism set a great model for me to follow.

To my thesis committee members, Associate Professor Dr. Hj. Ali Mamat and Associate Professor Dr. Hj. Md. Nasir Sulaiman, I would like to express appreciation for their insightful comments, questions, criticisms, and suggestions on the work. Their critical appraisals of my papers and presentations are extremely valuable for the improvement in my thinking also their kindness and willingness to help is unforgettable. Their great patience and continual encouragement over the past several years have been indispensable for the completion of this work.

In addition, I would like to express my gratitude to Associate Professor Dr. Abdul Azim, for his critical appraisals and constructive comments of my presentations are extremely valuable for the improvement in my research, also my special thanks to Associate Professor Dr. Ramlan Mahmod, he taught me how to do research.

This research is partially supported by an IRPA fund number (04-02-04-0797-EA001), which is sponsored by Malaysian Ministry of Science, Technology, and Innovation. Thanks to the Universiti Putra Malaysia and Malaysian Government for the support.

I would like to thank many people I have met during my stay in Malaysia for their help, enjoyable discussions and some good times.

Finally, I would like to express my love and deepest thanks to my parents, brothers and sisters, who have loved and supported me throughout my life. And also to my wife and my son, they made my life here in Malaysia enjoyable and memorable.

BELAL ZAQAIBEH

June 2006

I certify that an Examination Committee has met on 6th June 2006 to conduct the final examination of Belal Mohammad Zaqaibeh on his Doctor of Philosophy thesis entitled "Development of an Assertion Model of Integrity Constraints in Object-Oriented Databases" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Hjh. Fatimah Dato' Ahmad, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Abdul Azim Abdul Ghani, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Abdul Razak Hamdan, PhD
Professor
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
(External Examiner)

Mustafa Mat Deris, PhD
Professor
Faculty of Science and Technology
College University Since and Technology Malaysia
(External Examiner)

HASANAH MOHD GHAZALI, PhD
Professor / Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 15 JUN 2006

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee are as follows:

Hamidah Ibrahim, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Hj. Ali Mamat, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Hj. Md. Nasir Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

AINI IDERIS, PhD
Professor/Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 13 JUL 2006

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations, which have been duly acknowledged. I also declare that it has not been previously or concurrently for any other degree at UPM or other institutions.

BELAL M. ZAQAIBEH

Date: 14/6/2006

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADT	Abstract Data Type
AC	Antecedent Constraint
AID	Attribute ID
ALICE	Assertion Language for Integrity Constraint Expression
AMIC	Assertion Model of Integrity Constraints
AOIM	Application-Oriented Integrity Maintenance
ARIEL	ARray-orIEnted Language
ARL	A constraint Rule Language
BLOBS	Binary Large Objects
BPP	Backward Propagation Problem
CA	Constraint Analyzer
CAD	Computer-Aided Design
CAM	Computer-Aided Manufacturing
CASE	Computer-Aided Software Engineering
CC	Constraint Checker
CCG	Constraint Code Generator
CID	Class ID
CIM	Centralized Integrity Maintenance
CM	Constraint Maintenance
CTM	Compile-Time Model
CO	Constraint Optimizer
CP	Constraint Parser
DAG	Directed Acyclic Graph
DBMS	Database Management System
DE	Dependency Evaluation

DID	Domain ID
DM	Detection Method
EBNF	Extended Backus-Naur Form
EC	Engineering Changes
ECA	Event Condition Action
EH	Error Handler
ER	Entity Relational
GCS	Greatest Consistent Specialization
GIS	Geographic Information System
HiPAC	High Performance ACtive
IMPR	Integrity Maintenance Production Rule
IRules	Integration Rules
LAR	Limited Ambiguity Rule
LIC	Logical Integrity Constraint
OALIC	Object Assertion Language for Integrity Constraints
OCL	Object Constraint Language
ODE	Object Data and Environment
ODMG	Object Data Management Group
OID	Object Identifier
OIS	Office Information System
OM	Optimization Method
OMD	Object Meta Data
OMD^{CK}	OMD Constraint Knowledge
OMD^{CO}	OMD Constraints Optimization
OMD^{KB}	OMD Knowledge Base
OMG	Object Management Group

OODB	Object-Oriented Database
OODBMS	Object-Oriented Database Management System
OODM	Object-Oriented Data Model
RAID	Related Attribute ID
RCID	Related Class ID
RDB	Relational Database
RDBMS	Relational Database Management System
RDM	Relational Data Model
RMS	Rule Management System
RTM	Run-Time Model
SC	Supplement Constraint
SIC	Structural Integrity Constraint
SIS	Semantic Integrity Subsystem
SQL	Structured Query Language
TQL	Terminology Query Language
UA	Update analyzer
UC	Update Checker
UDC	User-Defined Constraint
UDT	User-defined Data Types
UE	Update Enforcer
UI	User Interface
UM	Update Maintenance
UML	Unified Modeling Language

CHAPTER 1

INTRODUCTION

1.1 Background

Integrity constraints refer to the expression of integrity validity and do not include the enforcement or the maintenance part. The term integrity covers consistency (data is well organized in accordance with the requirements of a data model) and validity (all invalid data is excluded from the database).

The proper handling of integrity constraints is essential to any data storage and management. Handling integrity constraints is an essential premise to managing semantically rich data (Formica, 2002; Rao, 1994). In Object-Oriented Databases (OODBs), checking the integrity constraints is a fundamental problem in the database design (Formica, 2002). The automated verification of constraints and their enforcement provided by current OODB Management Systems (OODBMSs) is limited (Formica, 2002; Eick and Werstein, 1993) due to the user participation is required.

Maintaining constraints that are scattered in applications is called Application-Oriented Integrity Maintenance (AOIM) (Do *et al*, 1997; Eick and Werstein, 1993). Centralizing the management of integrity constraints by extending database systems to have a dedicated component for constraint enforcement is called Centralized Integrity Maintenance (CIM) (Do *et al*, 1997; Eick and Werstein, 1993; Urban and Desiderio, 1992).

OODBMSs do not have adequate support for certain types of constraints especially the ones defined in a class composition and inherence hierarchies (Bagui, 2003; Formica, 2002; Do *et al*, 2002, 1997; Choi *et al*, 1997; Junmang, 1997; Do and Choi, 1994). The integrity constraints must be maintained in the backward direction along the class composition and inheritance hierarchies as well as in the forward direction.

The class composition hierarchy is represented by IS-PART-OF relationship, and class inheritance hierarchy is represented by IS-A relationship (Graham, 2001; Brown, 2001; David and Embley, 1998). The Object-Oriented Data Model (OODM) can support three types of relationships between classes, which are:

- Composition hierarchy (logical or physical composition) is a relationship between two classes where the instances of one class are in someway attributes, methods, and constraints of the other.
- Inheritance hierarchy (single or multiple inheritance) is a relationship between superclasses and subclasses. A superclass may have any number of subclasses, which subclasses inherit attributes and methods of superclass. This means all global attributes, methods, and constraints